DOCUMENT RESUME

ED 354 499 CS 011 208

AUTHOR Brisbois, Judith E.

TITLE Do First Language Writing and Second Language Reading Equal Second Language Reading Comprehension? An Assessment Dilemma.

PUB DATE Dec 92

NOTE 32p.; Paper presented at the Annual Meeting of the National Reading Conference (42nd, San Antonio, TX, December 2-5, 1992).

PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.

DESCRIPTORS College Students; French; Higher Education; Protocol Analysis; *Reading Comprehension; Reading Research; *Reading Writing Relationship; Regression (Statistics); *Second Language Learning; Test Interpretation; Test Validity; *Writing Ability

IDENTIFIERS Air Force Academy CO; Nelson Denny Reading Tests

ABSTRACT

A study focused on enhancing the validity of foreign language (L2) reading comprehension assessment measures. Subjects included 131 cadets (88 beginners and 43 non-beginners) enrolled in French courses at the United States Air Force Academy in Colorado Springs, Colorado. First language (L1) reading comprehension was measured using the comprehension portion of the Nelson-Denny Reading Test (NDRT) and the mean percentage score on three immediate timed written recall protocols. L1 writing ability, L2 vocabulary knowledge, L2 grammatical skill, and L2 reading comprehension also were measured. Six programs of hierarchical multiple regression were employed. Results indicated that: (1) means were higher for non-beginners than for beginners on all variables; (2) integrative tests such as the recall protocol provided a more sensitive measure of reading comprehension; (3) as reading skill increased, the more the sensitivity and enhanced discriminatory power of integrative tests came to the fore; (4) L2 vocabulary scores contributed more to the L2 reading scores than did the scores for any other independent variable for all subjects and for beginners versus non-beginners; and (5) unlike the regression models for beginners and non-beginners using NDRT comprehension scores, L1 reading skill (rather than L2 vocabulary knowledge) was the major contributor for both groups. Findings suggest that although the contribution of L1 writing ability to the scores on L2 reading tests appears to be slight, it should be taken into consideration when interpreting test scores. (Eight tables of data and 2 figures are included; 43 references are attached.) (RS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED354499

Do First Language Writing and Second Language Reading Equal Second Language Reading Comprehension? An Assessment Dilemma

by

Judith E. Brisbois

2637 Zephyr Dr.

Colorado Springs, CO 80920

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Judith E. Brisbois

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CS011208

BEST COPY AVAILABLE

Introduction

This study focuses on enhancing the validity of foreign language (L2) reading comprehension assessment measures. In recent years, increased discussion has ensued regarding the validity of L2 reading tests. One facet of the assessment debate centers on the decrease in construct validity that results when readers are required to write their responses on L2 reading measures in the L2. A number of studies have demonstrated that such tests reveal less evidence of comprehension than do those that call for responses in the native language (L1) (Hock & Poh, 1979; Lee, 1986; Shohamy, 1984). According to Weber (1991), when L2 readers respond "in their weaker second language, they may be limited in their ability to demonstrate their level of understanding" (p. 109). Thus, the practice of requiring responses in the L2 results in a confounding of the productive skill of writing and the receptive skill of reading.

Although L2 reading tests calling for L1 responses yield scores that more accurately represent reading comprehension (Bernhardt, 1991; Hock & Poh, 1979; Lee, 1986; Shohamy, 1984; Weber, 1991), the reception-production dilemma may still influence test scores. In this case, the fusion of L1 writing and L2 reading could result in scores representing a blend of both factors, thus decreasing test validity.

The purpose of the present study was to enhance L2 reading test validity by measuring the unique contributions of selected L1 and L2 literacy variables--L1 writing ability, L1 reading comprehension, L2 vocabulary knowledge, and L2 grammatical skill--to the scores on L2 reading comprehension tests. Additionally, it has investigated whether or not these contributions differ by course level.

Theoretical and Empirical Support

Test Theory

Test validity "is the most important consideration in test evaluation," according to a joint commission of the American Psychological Association, the American Educational Research Association, and the National Council of Mathematics Educators (in Bernhardt, 1991, p. 192). According to Messick (1988), the unifying force of test validity is "empirically grounded construct interpretation" (p. 35). Thus, because construct validity serves as the cornerstone of test validity, it is impossible for test validity to exist without "construct-related evidence" (Messick, 1988, p. 35). In order for a test to provide construct validity, it must "match as closely as possible what is known about a process at any point in time" (Bernhardt, 1991, p. 192).

An instrument that claims to measure the construct of L2 reading comprehension, therefore, must tap it in a manner that accurately reflects current theory- and research-based knowledge

of the L2 reading comprehension process. The accuracy of the interpretations and the appropriateness of the decisions that are based on the results of L2 reading comprehension measures hinge on the degree of construct validity that these tests provide.

The Melding of Literacy Transfer Theory and the Reading-Writing Relationship

Cummins's (1981) Interdependence Hypothesis works in conjunction with reading-writing relationship theory to buttress the proposition that L1 writing ability constitutes a portion of the scores on L2 reading assessment measures. The Interdependence Hypothesis asserts that literacy skills can transfer between languages, and maintains that experience with either the L1 or the L2 promotes development of the capacities underlying both languages.

An abundance of research on language transfer supports Cummins's (1981) claims. The studies of Groebel (1980), Carson, Carrell, Silberstein, Kroll, and Keuhn (1990), Goldman, Reyes, and Varnhagen (1984), Reyes (1987), and Wagner, Spratt, and Ezzaki (1989), among many others, indicate that reading skill transfers between languages, while the research of Canale, Frenette, and Bélanger (1988), Edelsky (1982), Jones and Tetroe (1987), Lay (1982), Mohan and Lo (1985), and Arndt (1987) point to the phenomenon of writing ability transfer.

The notion of a reading-writing relationship, like that of language transfer, has long been sustained by theory and research. Wittrock (1984) and Squire (1984) view reading skill and writing ability as employing the same cognitive processes. According to Wittrock (1984), both reading and writing involve "generative cognitive processes" (p. 600) that allow the learner to construct meaning by relating the text to personal experience. Squire (1984) evinces agreement with Wittrock by stating that composing and comprehending are "two sides of the same basic process" (p. 581) of constructing and reconstructing ideas. Tierney and Pearson (1983) further develop the notion of shared cognitive processes in reading and writing through their composing model of reading, which asserts that both reading and writing are essentially composing processes (in Tierney & Shanahan, 1991).

In the research arena, numerous studies clearly point to a relationship between reading and writing skills, indicating the existence of overlapping cognitive processes. The studies of Carson et al. (1990), Juel, Griffith, and Gough (1986), Langer (1984), Schewe and Froese (1986), Shanahan (1984), and Shanahan and Lomax (1986, 1988), among many others, suggest the likelihood of shared knowledge in reading and writing.

In summary, theoretical and empirical evidence point to the occurrence of language transfer and to the existence of a relationship between reading skill and writing ability. Evidence

of a reading-writing relationship, coupled with that of the transfer of skills between languages, leads to the hypothesis that L1 literacy abilities contribute to the scores on L2 reading tests.

The L2 Reading Comprehension Process

Bernhardt's (1986) Constructivist Reading Model contributes a valuable theoretical foundation to this study because it takes into account the roles in the L2 reading comprehension process of two of the literacy variables that were explored: L2 vocabulary knowledge and L2 grammatical skill. This six-part model integrates both text-based and extratext-based components. The text-based elements include phonemic/graphemic features, syntactic feature recognition (grammatical ability), and word recognition (vocabulary knowledge), while the extratext-based features comprise intratextual perceptions, prior knowledge, and metacognition. Because the model is non-linear, the six elements are free to interact and to meld together in varying combinations in order to forge the construct of L2 reading comprehension (Bernhardt, 1986).

This model is particularly relevant to the present study because the unique contributions of its elements to L2 reading comprehension remain unspecified. Such flexibility is crucial because the contributions may differ by course level, or by a host of other variables (Everson, 1986). The present study has aided

in the testing of this model because it has explored the extent to which word recognition and syntactic feature recognition contribute to L2 reading comprehension scores.

Methodology

Subjects

The subjects included 131 cadets (88 Beginners and 43 Non-Beginners) enrolled in French courses at the United States Air Force Academy (USAFA) in Colorado Springs, CO. The Beginners were enrolled in Basic French. The Non-Beginner group consisted of 21 intermediate-level and 22 advanced-level subjects.

Instrumentation

Independent Variables

L1 Reading Comprehension.

Two separate testing instruments were used to assess L1 reading comprehension: (a) the comprehension portion of the Nelson-Denny Reading Test (NDRT), Form E (Brown, Bennett, & Hanna, 1981), and (b) the mean percentage score on three immediate timed written recall protocols. The comprehension portion of the NDRT consists of eight reading passages, followed by a total of 36 multiple choice questions.

The three recall passages, selected from the Encyclopaedia Britannica (1984), were each between 200 and 250 words in length, and historical in nature. Two trained native English-speaking raters scored the protocols using the Johnson system (Johnson,

1970). Interrater reliability, measured using the Pearson product-moment correlation procedure, was .94.

L1 Writing Ability.

This variable was measured using the combined holistic scores on the same three recall protocols that were employed to assess L1 reading comprehension. In order to arrive at the scores for writing ability, four trained, experienced raters used the Test of Written English Scoring Guide (Educational Testing Service, 1990). Overall interrater reliability, as measured by the Spearman-Brown Prophecy Formula (Henning, 1987), was .95.

L2 Vocabulary Knowledge.

A supply-definition test was used to measure L2 vocabulary knowledge. Two experienced French language instructors from the Air Force Academy selected the 50 words appearing in the French reading passages (See Dependent Variable: L2 Reading Comprehension) that the beginning-level subjects would find the most difficult to define. (The beginning level was used as the base line.) For each French word, the subjects were required to write, in English, either the definition or the English equivalent. Two French language experts scored each answer on a two-point scale consisting of the following: (a) zero points for an incorrect response, (b) one point for a partially correct response, and (c) two points for a correct response. Interrater

reliability, as measured by the Pearson product-moment correlation coefficient, was .99.

L2 Grammatical Skill.

Scores for French grammatical skill were obtained from the grammar portion of USAFA's French placement test. This section of the test features multiple choice and cloze items.

Dependent Variable: L2 Reading Comprehension

This variable was measured using the mean percentage scores on three immediate free written recall protocols. The three French passages, each between 200 and 250 words in length and historical in nature, were selected from the Encyclopaedia Universalis (1968), the French counterpart to the Encyclopaedia Britannica. Texts of the same genre as those used in the L1 recall procedure were selected in order to maintain consistency in discourse type. Two French language experts scored the protocols using the Johnson system (Johnson, 1970). Interrater reliability, measured using the Pearson product-moment correlation coefficient, was .94.

Procedures and Data Collection

Procedures and data collection will be explained in the order in which the tests were administered. Data collection took place during the fall semester of 1991.

NDRT.

USAFA personnel administer the NDRT to all cadets entering the Academy as part of the admissions procedure. Scores were obtained through the Department of English, USAFA.

L2 Grammatical Skill.

Two months before the recall protocol data were collected, the French placement test was administered to the beginning-level subjects by members of the USAFA French Division as part of the Academy's admissions procedure. The researcher administered the grammar portion to the intermediate- and advanced-level subjects during class time three weeks before the recall protocol data were collected.

L2 Vocabulary Knowledge.

Two weeks before the recall protocol data were collected, the USAFA French instructors assigned the vocabulary test to be completed outside of class. The subjects were told that they would be under the provisions of the Cadet Honor Code not to use dictionaries or to seek help in defining the words. They were allowed a maximum of 50 minutes (one minute per word) to complete the test. The course instructors collected the tests during the following class period.

Recall Protocols.

The recall protocols were collected over a four-day period. Because the USAFA-wide class schedule is composed of two-day

cycles, some of the French classes meet the first day of the cycle while the others meet the second. The subjects completed the recall protocols based on L1 passages during the first two-day cycle and those based on L2 texts during the second cycle.

Before the experiment, the researcher used the class rosters to determine the number of subjects at each language level, and then assembled the appropriate number of research packages. The texts were arranged in differing orders in order to limit the effects of fatigue. Subjects were selected through the use of a table of random numbers (Snedecor & Cochran, 1967) to receive the various arrangements.

When the L1 recall protocols were administered, the subjects were told that the research packets contained three exercises and that they would have five minutes to complete each one. The time limit was imposed in order to decrease the possibility of the subjects memorizing the passages. Additionally, the subjects were advised that upon opening the packet to the first passage, they would have a total of five minutes to read it and, without referring back to it, to write, in English, everything they could remember about it. After the first exercise was completed, the same procedure was used for the second and third.

The same procedures were followed when the recall protocols based on L2 reading passages were administered, with two exceptions. First, the subjects had the entire 50-minute class

period to read and recall (in English) the French texts. Second, because the subjects were working at their own pace, they were reminded to complete the readings in the order given and that once they completed one recall exercise and began another, they could not go back. After each subject had completed the three protocols, his or her packet was collected.

Data Analysis

Six programs of hierarchical multiple regression were employed. Individual programs were run for All Subjects, for Beginners, and for Non-Beginners. Additionally, because L1 reading comprehension was measured using two different testing instruments, separate programs were performed for each group, first using the NDRT comprehension scores, and then using the recall protocol scores.

In all of the hierarchical regression programs used, the scores for the independent variables were entered into the statistical model in the following order: (a) L1 reading, (b) L1 writing, (c) L2 vocabulary, and (d) L2 grammar. The L1 variables were introduced first because the subjects, all native English speakers, were considerably more experienced with the English language than they were with the L2. L1 reading data were inserted before the L1 writing scores because, according to Shanahan and Lomax (1986), "more reading information is used in writing than vice versa" (p. 122). L2 vocabulary scores were

entered before those for L2 grammar because L2 vocabulary skill consistently contributed more to L2 reading comprehension scores than did L2 grammatical knowledge in the stepwise multiple regression analyses performed by Henning (1975) and Koda (1990).

Results and Discussion

Means and Standard Deviations

Three principal findings have emerged from an analysis of the means, standard deviations, and minimum and maximum scores. (See Tables 1 and 2.) First, the means are higher for Non-Beginners than for Beginners on all variables. This dissimilarity for the L2 variables is consistent with the course curriculum in which the subjects are increasingly exposed to elements of L2 literacy. Unexpected however, were the differences for the L1 variables. These data suggest that language students opting to continue their L2 studies comprise a select group of highly literate individuals.

Second, for both Beginners and Non-Beginners, the mean NDRT comprehension scores tend to be concentrated closer to the maximum possible score than are the mean recall scores. Because of the cluster of high NDRT comprehension scores, the distributions for All Subjects, Beginners, and Non-Beginners were profoundly negatively skewed. (See Figure 1.) In contrast, the distributions of the recall scores for the three subject groups were nearly normal. (See Figure 2.) Unlike the near-normal distributions of recall scores, the negatively skewed

Table 1: Means and Standard Deviations for Literacy Variables for Beginners

	N	Mean	SD	Min.	Max.
1. L1 Reading					
a. NDRT	87	25.87	5.97	8	34
b. Protocols	87	36.44	8.71	18	63
2. L1 Writing	87	10.31	2.24	5	15
3. L2 Vocabulary	88	9.76	6.34	0	32
4. L2 Grammar	86	22.07	6.13	1	39
5. L2 Reading	88	19.73	6.53	7	39

Table 2: Means and Standard Deviations for Literacy Variables for Non-Beginners

	N	Mean	SD	Min.	Max.
1. L1 Reading					
a. NDRT	39	29.64	3.44	21	35
b. Protocols	42	44.19	9.75	25	62
2. L1 Writing	41	12.00	3.22	5	18
3. L2 Vocabulary	42	39.50	14.91	10	81
4. L2 Grammar	43	42.56	5.65	29	54
5. L2 Reading	43	38.53	10.22	21	62

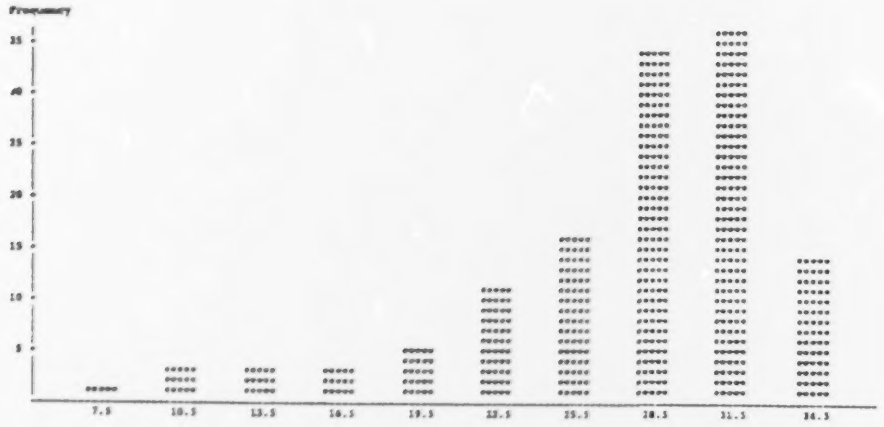

Figure 1. Frequencies of NDRT Comprehension Scores for All Subjects

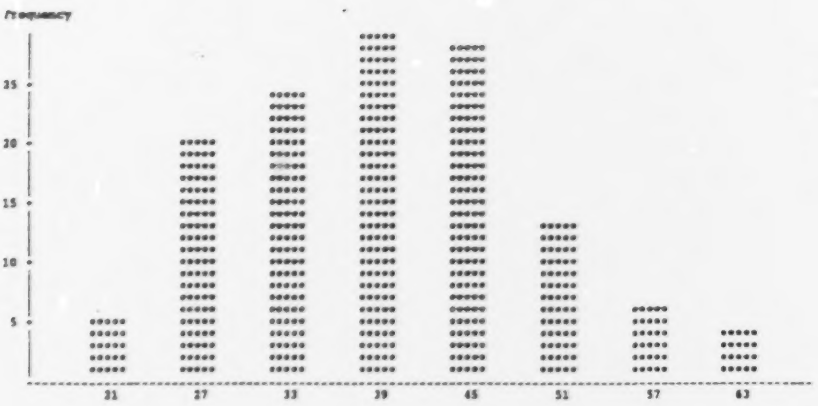

Figure 2. Frequencies of L1 Recall Protocol Scores for All Subjects

distributions of NDRT comprehension scores exhibit minimal variability, pointing to limitations in the test's discriminatory power. In other words, integrative tests such as the recall protocol clearly provide a more sensitive measure of reading comprehension.

Third, the standard deviation of the NDRT comprehension scores is smaller for Non-Beginners than it is for Beginners, while the opposite is true for the standard deviation of the recall scores. In other words, the Non-Beginners' NDRT comprehension scores are less variable than those of the Beginners, while the Non-Beginners' recall scores are more variable. Thus, as reading skill increases, the greater sensitivity and enhanced discriminatory power of integrative tests come to the fore.

Hierarchical Regression Analyses

All Subjects

Regardless of the instrument used to assess L1 reading comprehension, the independent variables contributed in the following order of importance: (a) L2 vocabulary, (b) L1 reading, (c) L1 writing, and (d) L2 grammar. (See Tables 3 and 4.) The finding that the L2 vocabulary scores contributed more to the L2 reading scores than did the scores for any other independent variable, particularly grammatical skill, is consistent with the

Table 3: Summary of Hierarchical Regression Analysis for All Subjects using NDRT Comprehension Scores

Steps/Variables	R^2 Increment	Cum. R^2	F	p
1. L1 Reading	.1412	.1412	20.38	.0001
	(126)			
2. L1 Writing	.0703	.2115	10.70	.0500
	(123)			
3. L2 Vocabulary	.3568	.5683	58.27	.0001
	(123)			
4. L2 Grammar	.0304	.5987	8.78	.0500
	(121)			

Note. Sample size is indicated in parentheses.

Table 4: Summary of Hierarchical Regression Analysis for All Subjects using L1 Protocol Scores

Steps/Variables	R^2 Increment	Cum. R^2	F	p
1. L1 Reading	.2749	.2749	47.77	.0001
	(128)			
2. L1 Writing	.0384	.3133	6.93	.0200
	(127)			
3. L2 Vocabulary	.3091	.6224	99.84	.0001
	(126)			
4. L2 Grammar	.0319	.6543	10.97	.0050
	(124)			

Note. Sample size is indicated in parentheses.

results of the stepwise regression studies performed by Henning (1976) and Koda (1990).

Although L2 vocabulary knowledge was the primary contributor, the L1 variables offered a substantial cumulative contribution in both models (21.15% using NDRT comprehension scores and 31.33% using L1 recall scores). This finding not only bolsters Cummins's (1981) Interdependence Hypothesis, but corroborates the results of previous research pointing to the transfer phenomenon between L1 and L2 reading skills (Carson et al., 1990; Groebel, 1980; Reyes, 1987; Wagner et al., 1989) and to those indicating that reading ability and writing skill transfer to one another (Carson et al., 1990; Juel et al., 1986; Langer, 1984; Schewe & Froese, 1986).

The results for All Subjects also indicate that some of the elements in Bernhardt's (1986) Constructivist reading Model contribute more to the reading process than do others. According to the data, L2 vocabulary knowledge contributes more to L2 reading comprehension than does L2 grammatical skill. Additionally, prior knowledge in the form of L1 literacy abilities contributes a great deal to L2 reading. Evidently, language students depend upon their L1 skills in order to make sense of the L2.

The most intriguing difference between the results of the two regression models is the dissimilarity in the contributions of the L1 reading scores when assessed using the NDRT comprehension test

versus recall protocols. When recall scores were used, L1 reading contributed nearly twice the variance (27.49%) as it did when NDRT comprehension scores were used (14.12%). This situation came about because the more normally distributed recall scores (See Figure 2) had more variance to contribute. The fact that the overall picture of reading comprehension changes depending on the testing instrument used to measure it underscores the need for multiple measures.

Additionally, a greater percentage of the overall variance was accounted for in the regression model using recall scores (65.43%) versus the one featuring NDRT comprehension scores (59.87%). This difference again relates to the greater variability in the recall scores. The more sensitive the assessment measure is, the more accurate the information will be about the components of the L2 reading comprehension process.

Beginners versus Non-Beginners using NDRT Comprehension Scores

For Beginners, the order of importance for the independent variables was: (a) L2 vocabulary (8.50%), (b) L1 reading (7.06%), (c) L1 writing (3.47%), and (d) L2 grammar (1.24%). (See Table 5.) For Non-Beginners, the order was: (a) L2 vocabulary (8.67%), (b) L1 writing (5.85%), (c) L1 reading (4.42%), and (d) L2 grammar (1.49%). (See Table 6.)

Consistent with the results for All Subjects, L2 vocabulary knowledge contributed more to the L2 reading scores than did any

Table 5: Summary of Hierarchical Regression Analysis for Beginners Using NDRT Comprehension Scores

Steps/Variables	R^2 Increment	Cum. R^2	F	p
1. L1 Reading	.0706	.0706	6.46	.0129
	(87)			
2. L1 Writing	.0347	.1053	3.22	.0600
	(86)			
3. L2 Vocabulary	.0850	.1903	8.61	.0043
	(86)			
4. L2 Grammar	.0124	.2027	1.23	.2000
	(84)			

Note. Sample size is indicated in parentheses.

Table 6: Summary of Hierarchical Regression Analysis for Non-Beginners Using NDRT Comprehension Scores

Steps/Variables	R^2 Increment	Cum. R^2	F	p
1. L1 Reading	.0442	.0442	1.71	.1992
	(39)			
2. L1 Writing	.0585	.1027	2.21	.1500
	(37)			
3. L2 Vocabulary	.0867	.1894	3.53	.0691
	(37)			
4. L2 Grammar	.0149	.2043	.60	.4444
	(37)			

Note. Sample size is indicated in parentheses.

of the other independent variables, especially grammatical skill. This finding coincides with the results of the regression studies performed by Henning (1975) and Koda (1990). Additionally, the cumulative contribution of the L1 variables was greater than that of the L2 variables for both groups, further reinforcing Cummins's (1981) Interdependence Hypothesis.

Furthermore, these data aid in expanding Bernhardt's (1986) Constructivist Reading Model by exploring the contributions of several of its components by course level. Vocabulary knowledge was shown to be a crucial element in the reading comprehension process, slightly more so for Non-Beginners (8.67%) than for Beginners (8.50%). Additionally, the cumulative contribution of the L2 variables was greater for Non-Beginners than for Beginners, while the opposite was true for that of the L1 variables. Evidently, language students rely on prior knowledge in the form of L1 skills until their L2 abilities develop and thus become increasingly useful in the reading comprehension process.

Both models resulted in nearly equal portions of the total variance unaccounted for: 79.73% for Beginners and 79.57% for Non-Beginners. Clearly, L2 readers draw upon components of the reading process that were not included in the present set of independent variables.

Beginners versus Non-Beginners using Recall Protocol Scores

For Beginners, the independent variables contributed as follows: (a) L1 reading (10.58%), (b) L2 vocabulary (8.38%), L1 writing (3.00%), and (d) L2 grammar (1.85%). (See Table 7.) The order of importance for Non-Beginners was: (a) L1 reading (23.96%), (b) L2 vocabulary (5.40%), (c) L2 grammar (2.69%), and (d) L1 writing (0.43%). (See Table 8.)

Unlike the regression models for Beginners and Non-Beginners using NDRT comprehension scores, L1 reading skill (rather than L2 vocabulary knowledge) was the major contributor for both groups. This difference resulted from the increased variability in the recall scores, particularly for Non-Beginners. The fact that L1 reading ability contributed more than twice the variance for Non-Beginners (23.96%) as it did for Beginners (10.58%) indicates that the sensitivity of the recall protocol increases along with language level.

Consistent with the results of the programs for Beginners and Non-Beginners using NDRT comprehension scores, the regression models using recall scores show a higher cumulative contribution for the L1 variables than for the L2 variables. Once again, this finding supports Cummins's (1981) Interdependence Hypothesis. Furthermore, and coinciding with the results of the other regression models in this study, L2 vocabulary contributed more to the L2 reading scores than did L2 grammatical skill.

Table 7: Summary of Hierarchical Regression Analysis for Beginners Using L1 Protocol Scores

Steps/Variables	R^2 Increment	Cum. R^2	F	p
1. L1 Reading	.1058	.1058	10.06	.0021
	(87)			
2. L1 Writing	.0300	.1358	2.91	.0916
	(87)			
3. L2 Vocabulary	.0838	.2196	8.91	.0037
	(87)			
4. L2 Grammar	.0185	.2381	1.95	.2000
	(85)			

Note. Sample size is indicated in parentheses.

Table 8: Summary of Hierarchical Regression Analysis for Non-Beginners Using L1 Protocol Scores

Steps/Variables	R^2 Increment	Cum. R^2	F	p
1. L1 Reading	.2396	.2396	12.60	.0010
	(42)			
2. L1 Writing	.0043	.2439	.21	.5000
	(41)			
3. L2 Vocabulary	.0540	.2979	2.74	.1500
	(40)			
4. L2 Grammar	.0269	.3248	1.38	.2485
	(40)			

Note. Sample size is indicated in parentheses.

The regression models using recall scores show 76.19% of the overall variance unaccounted for in the data for Beginners and 67.52% unaccounted for in the Non-Beginner group. Again, it is evident that L2 readers draw on components of the L2 reading process that were not included in this study.

Implications

Assessment

In the area of assessment, this study has opened a window on the dilemma regarding L2 reading test validity. The results of the models for All Subjects show the effect of L1 writing ability to be small but significant. Thus, although the contribution of L1 writing ability to the scores on L2 reading tests appears to be slight, it should be taken into consideration when interpreting test scores.

These data will lead to enhanced test development and to improvements in diagnostic and placement testing. In terms of test development, the knowledge that L1 writing ability is a confounding factor, albeit a minor one, points to the need for testing methods that will decrease or eliminate its effects. For diagnostic testing, perhaps L1 literacy abilities need to be tested in order to gain insight into learner difficulties in L2 study. Regarding placement testing, L1 literacy abilities, especially reading skill, have been identified as important predictors of L2 literacy ability. L1 literacy skills, therefore,

should be considered when placing students into language course levels. Additionally, because different results emanate from different testing instruments, the use of multiple measures is necessary in order to arrive at an accurate picture of what has been comprehended.

Pedagogy

In the pedagogical domain, this study attests to the critical nature of vocabulary knowledge to L2 reading comprehension. Instructors and curriculum developers should continue to stress vocabulary skill as a crucial element of L2 study.

Additionally, the data underline the importance of capitalizing on students' already developed L1 knowledge. Instructors should draw attention to parallels and contrasts between the students' L1 and the L2, and encourage learners to transfer their skills between languages.

In programs in which the learners are already literate in their native languages, the importance of L1 literacy abilities to the development of L2 literacy should be increasingly reflected in L2 curriculum development. In such programs, beginning-level courses should aid students in going from the known to the unknown by emphasizing the use of L1 knowledge in L2 study. Follow-on courses should guide learners to increased use of their developing L2 knowledge. Furthermore, students experiencing difficulties in the L2 may benefit from further L1 instruction.

Research

Research is needed that will validate and expand upon the results of the present study. Replications should include subjects of diverse ages and language backgrounds. Additionally, because gross measures of reading and writing were used in the present study, replications are needed in which finer assessment measures will be employed. Alternate data analysis techniques could also be used. The independent variables in the present study, for example, should be entered into programs of hierarchical multiple regression in varying orders. Qualitative methods could be employed. Furthermore, because the variance in the L2 reading scores is still not completely accounted for, additional independent variables need to be considered.

In addition, integrative testing methods, such as the recall protocol procedure, need to gain wider acceptance. Not only is the recall protocol more sensitive than discrete-point tests, but its sensitivity becomes more pronounced as reading proficiency increases.

Most importantly, the results of this study accentuate the need for multiple measures. Multiple testing instruments will lead to a clearer view of the constructs being measured, and hence to more astute research insights and to more perceptive theory building.

References

Alexander, L., & James, T. (1987). The nation's report card: Improving the assessment of student achievement: Report of the Study Group. Cambridge, MA: National Academy of Education.

Arndt, V. (1987). Six writers in search of text: A protocol-based study of L1 and L2 writing. English Language Teaching Journal, 41(4), 257-267.

Bélanger, J. (1987). Theory and research into reading and writing connections: A critical review. Reading-Canada-Lecture, 5(1), 10-18.

Bernhardt, E. (1986). Reading in a foreign language. In B. Wing (Ed.), Listening, reading, writing: Analysis and application (pp. 93-115). Middlebury, VT: Northeast Conference on the Teaching of Foreign Languages.

Bernhardt, E. (1991). Reading development in a second language. Norwood, NJ: Ablex Publishing Corporation.

Brown, J., Bennett, J., & Hanna, G. (1981). The Nelson-Denny Reading Test. Chicago: Riverside Publishing Co.

Canale, M., Frenette, N., & Bélanger, M. (1988). Evaluation of minority student writing in first and second language. In J. Fine (Ed.), Second language discourse: A textbook of current research. Norwood, NJ: Ablex Publishing Corporation.

Carson, J., Carrell, P., Silberstein, S., Kroll, B., & Kuehn, P. (1990). Reading-writing relationships in first and second language. TESOL Quarterly, 24(2), 245-266.

Cummins, J. (1981). Bilingualism and minority-language children: Language and literacy series. Toronto: Ontario Institute for Studies in Education. (ERIC Document Reproduction Service No. ED 215 557)

Edelsky, C. (1982). Writing in a bilingual program: The relation of L1 and L2 texts. TESOL Quarterly, 16(2), 211-228.

Educational Testing Service (1990). Test of Written English (TWE) Scoring Guide. Princeton, NJ: Educational Testing Service.

Encyclopedia Britannica (1984). Chicago, IL: Encyclopedia Britannica, Inc.

Encyclopedia Universalis (1968). Clichy, France: Paul Dupont.

Everson, M. (1986). The effect of word-unit spacing upon the reading strategies of native and non-native readers of Chinese: An eye-tracking study. Unpublished doctoral dissertation. The Ohio State University, Columbus, OH.

Goldman, S., Reyes, M., & Varnhagen, C. (1984). Understanding fables in first and second languages. NABE Journal, 8(2), 35-66.

Groebel, L. (1980). A comparison of students' reading comprehension in their native language with their reading comprehension in the target language. English Language Teaching Journal, 35, 54-59.

Henning, G. (1975). Measuring foreign language reading comprehension. Language Learning, 25(1), 109-114.

Henning, G. (1987). A guide to language testing. Cambridge, MA: Newbury House.

Hock, T., & Poh, L. (1979). The performance of a group of Malay-medium students in an English reading comprehension test. Regional Language Centre Journal, 10(1), p. 81-87.

Johnson, R. (1970). Recall of prose as a function of the structural importance of linguistic units. Journal of Verbal Learning and Verbal Behavior, 9, 12-20.

Jones, S., & Tetroe, J. (1987). Composing in a second language. In A. Matsuhashi (Ed.), Writing in real time: Modeling production process. Norwood, NJ: Ablex Publishing Corporation.

Juel, C., Griffith, P., & Gough, P. (1986), Acquisition of literacy: A longitudinal study of the children in first and second grade. Journal of Educational Psychology, 78(4), 243-255.

Koda, K. (1990). Factors affecting second language text comprehension. J. Zutell and S. McCormick (Eds.), Thirty-ninth Yearbook of the National Reading Conference (pp. 419-427). Chicago, IL: The National Reading Conference, Inc.

Langer, J. (1984). The effects of available information on responses to school writing tasks. Research in the Teaching of English, 18(1), 27-44.

Lay, N. (1982). Composing processes of adult ESL learners: A case study. TESOL Quarterly, 16(3), 406.

Lee, J. (1986). On the use of the recall task to measure L2 reading comprehension. Studies in second language acquisition, 8, 201-212.

Messick, S. (1988). The once and future issues of validity: Assessing the meaning and consequences of measurement. In H. Wainer & H. Braun (Eds.), Test validity (pp. 33-45). Hillsdale, NJ: Lawrence Erlbaum.

Mohan, B., & Lo, W. (1985). Academic writing and Chinese students: Transfer and developmental factors. TESOL Quarterly, 19(3), 515-534.

Reyes, M. (1987). Comprehension of content area passages: A study of Spanish/English readers in third and fourth grade. In S. Goldman and H. Trueba (Eds.), Becoming literate in English as a second language. Norwood, NJ: Ablex Publishing Corporation.

Schewe, A., & Froese, V. (1986, December). Relating reading and writing via comprehension, quality, and structure. Paper presented at the annual meeting of the National Research Conference, Austin, TX. (ERIC Document Reproduction Service No. ED 281 141)

Shanahan, T. (1984). Nature of the reading-writing relationship: An exploratory multivariate analysis. Journal of Educational Psychology, 76(3), 466-477.

Shanahan, T., & Lomax, R. (1986). An analysis and comparison of theoretical models of the reading-writing relationship. Journal of Educational Psychology, 78(2), 116-123.

Shanahan, T., & Lomax, R. (1988). A developmental comparison of three theoretical models of the reading-writing relationship. Research in the Teaching of English, 22(2), 196-212.

Shohamy, E. (1984). Does the testing method make a difference? The case of reading comprehension. Language Testing, 1, 147-170.

Snedecor, G, & Cochran, W. (1967). Statistical Methods. Ames, IA: Iowa State University Press.

Squire, J. (1984). Composing and comprehending: Two sides of the same basic processes. In J. Jensen (Ed.), Composing and comprehending (pp. 23-31). Urbana, IL: National Council of Teachers of English.

Tierney, R., & Pearson, D. (1983). Toward a composing model of reading. Language Arts, 60(5), 568-586.

Tierney, R., & Shanahan, T. (1991). Research on the reading-writing relationship: Interactions, transactions, and outcomes. In R. Barr, M. Kamil, P. Mosenthal, and P. Pearson (Eds.), Handbook of reading research. New York: Longman.

Wagner, D, Spratt, J., & Ezzaki, A. (1989). Does learning to read in a second language always put the child at a disadvantage? Some counterevidence from Morocco. Applied Psycholinguistics, 10, 31-48.

Weber, R. (1991). Linguistic diversity and reading in American society. In R. Barr, M. Kamil, P. Mosenthal, and P. Pearson (Eds.), Handbook of reading research (vol 2), 97-119. New York: Longman.

Wittrock, M. (1984). Writing and the teaching of reading. In J. Jensen (Ed.), Composing and comprehending (pp. 77-83). Urbana, IL: National Council of Teachers of English.